

SEQUENCE LISTING

(1). GENERAL INFORMATION:

(i) APPLICANT(S): Kaufman, Randal J.
Wasley, Louise

(ii) TITLE OF INVENTION: Method of Increasing Yield of
Mature Proteins

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: Massachusetts
(E) COUNTRY: United States of America
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb storage
(B) COMPUTER: IBM PS/2
(C) OPERATING SYSTEM: PC-DOS
(D) SOFTWARE: WordPerfect 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: not yet available
(B) FILING DATE: 26 November 1990
(C) CLASSIFICATION: not yet available

(vii) PRIOR APPLICATION DATA: not applicable

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION

(A) NAME: Ellen J. Kapinos, Esquire
(B) REGISTRATION NUMBER: 32,245
(C) REFERENCE/DOCKET NUMBER: GI 5181

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 876-1170
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) Topology: unknown

(ii) MOLECULE TYPE: partial human genomic DNA

- (A) DESCRIPTION: sequence encoding furin

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(v) PUBLICATION INFORMATION: A.M.W. van den Ouwehand et al,
Nucl. Acids. Res., 18:664 (1990)

(vi) SEQUENCE DESCRIPTION: SEQ ID NO:1

ATG GAG CTC AGG CCC TGG TTC 21
Met Glu Leu Arg Pro Trp Leu
1 5

CTA TGG GTC GTA CCA CCA ACA GGA ACC TTG GTC CTG CTA 60
Leu Trp Val Val Ala Ala Thr Gly Thr Leu Val Leu Leu
10 15 20

GCA GCT GAT GCT CAG GGC CAG AAG GTC TTC ACC AAC ACG 99
Ala Ala Asp Ala Gln Gly Gln Lys Val Phe Thr Asn Thr
25 30

TGG GCT GTG CGC ATC CCT GGA GGC CCA GCG GTG GCC AAC 138
Trp Ala Val Arg Ile Pro Gly Gly Pro Ala Val Ala Asn
35 40 45

AGT GTG GCA CGG AAG CAT GGG TTC CTC AAC CTG GGC CAG 177
Ser Val Ala Arg Lys His Gly Phe Leu Asn Leu Gly Gln
50 55

ATC TTC GGG GAC TAT TAC CAC TTC TGG CAT CGA GGA GTG 216
Ile Phe Gly Asp Tyr Tyr His Phe Trp His Arg Gly Val
60 65 70

ACG AAG CGG TCC CTG TCG CCT CAC CGC CCG CGG CAC AGC	255	
Thr Lys Arg Ser Leu Ser Pro His Arg Pro Arg His Ser		
75	80	85
CGG CTG CAG AGG GAG CCT CAA GTA CAG TGG CTG GAA CAG	294	
Arg Leu Gln Arg Glu Pro Gln Val Gln Trp Leu Glu Gln		
90	95	
CAG GTG GCA AAG CGA CGG ACT AAA CGG GAC GTG TAC CAG	333	
Gln Val Ala Lys Arg Arg Thr Lys Arg Asp Val Tyr Gln		
100	105	110
GAG CCC ACA GAC CCC AAG TTT CCT CAG CAG TGG TAC CTG	372	
Glu Pro Thr Asp Pro Lys Phe Pro Gln Gln Trp Tyr Leu		
115	120	
TCT GGT GTC ACT CAG CGG GAC CTG AAT GTG AAG GCG GCC	411	
Ser Gly Val Thr Gln Arg Asp Leu Asn Val Lys Ala Ala		
125	130	135
TGG GCG CAG GGC TAC ACA GGG CAC GGC ATT CTG GTC TCC	450	
Trp Ala Gln Gly Tyr Thr Gly His Gly Ile Val Val Ser		
140	145	150
ATT CTG GAC GAT GGC ATC GAG AAG AAC CAC CCC GAC TTG	489	
Ile Leu Asp Asp Gly Ile Glu Lys Asn His Pro Asp Leu		
155	160	
GCA GGC AAT TAT GAT CCT GGG GCC AGT TTT CAT GTC AAT	528	
Ala Gly Asn Tyr Asp Pro Gly Ala Ser Phe Asp Val Asn		
165	170	175
GAC CAG GAC CCT GAC CCC CAG CCT CGG TAC ACA CAG ATG	567	
Asp Gln Asp Pro Asp Pro Gln Pro Arg Tyr Thr Gln Met		
180	185	
AAT GAC AAC AGG CAC GGC ACA CGG TGT GCG GGG GAA GTG	606	
Asn Asp Asn Arg His Gly Thr Arg Cys Ala Gly Glu Val		
190	195	200
GCT GCC GTG GCC AAC AAC CGT GTC TGT GGT GTA GGT GTG	645	
Ala Ala Val Ala Asn Asn Gly Val Cys Gly Val Gly Val		
205	210	215
GCC TAC AAC GCC CGC ATT GGA GGG GTC CGC ATG CTG GAT	684	
Ala Tyr Asn Ala Arg Ile Gly Gly Val Arg Met Leu Asp		
220	225	
GGC GAG GTG ACA GAT GCA GTG GAG GCA CGC TCG CTG GGC	723	
Gly Glu Val Thr Asp Ala Val Glu Ala Arg Ser Leu Gly		
230	235	240
CTG AAC CCC AAC CAC ATC CAC ATC TAC AGT GCC AGC TGG	762	

Arg Leu Gln Arg Glu Pro Gln Val Gln Trp Leu Glu Gln
90 95

Gln Val Ala Lys Arg Arg Thr Lys Arg Asp Val Tyr Gln
100 105 110

Glu Pro Thr Asp Pro Lys Phe Pro Gln Gln Trp Tyr Leu
115 120

Ser Gly Val Thr Gln Arg Asp Leu Asn Val Lys Ala Ala
125 130 135

Trp Ala Gln Gly Tyr Thr Gly His Gly Ile Val Val Ser
140 145 150

Ile Leu Asp Asp Gly Ile Glu Lys Asn His Pro Asp Leu
155 160

Ala Gly Asn Tyr Asp Pro Gly Ala Ser Phe Asp Val Asn
165 170 175

Asp Gln Asp Pro Asp Pro Gln Pro Arg Tyr Thr Gln Met
180 185

Asn Asp Asn Arg His Gly Thr Arg Cys Ala Gly Glu Val
190 195 200

Ala Ala Val Ala Asn Asn Gly Val Cys Gly Val Gly Val
205 210 215

Ala Tyr Asn Ala Arg Ile Gly Gly Val Arg Met Leu Asp
220 225

Gly Glu Val Thr Asp Ala Val Glu Ala Arg Ser Leu Gly
230 235 240

Leu Asn Pro Asn His Ile His Ile Tyr Ser Ala Ser Trp
245 250

Gly Pro Glu Asp Asp Gly Lys Thr Val Asp Gly Pro Ala
255 260 265

Arg Leu Ala Glu Glu Ala Phe Phe Arg Gly Val Ser Gln
270 275 280

Gly Arg Gly Gly Leu Gly Ser Ile Phe Val Trp Ala Ser
285 290

Gly Asn Gly Gly Arg Glu His Asp Ser Cys Asn Cys Asp
295 300 305

Gly Tyr Thr Asn Ser Ile Tyr Thr Leu Ser Ile Ser Ser
310 315

Leu Asn Pro Asn His Ile His Ile Tyr Ser Ala Ser Trp
245 250

GGC CCC GAG GAT GAC GGC AAG ACA GTG GAT GGG CCA GCC 801
Gly Pro Glu Asp Asp Gly Lys Thr Val Asp Gly Pro Ala
255 260 265

CGG CTC GCC GAG GAG GCC TTC TTC CGT GGG CTT AGC CAG 840
Arg Leu Ala Glu Glu Ala Phe Phe Arg Gly Val Ser Gln
270 275 280

GGC CGA GGG GGG CTG GGC TCC ATC TTT GTC TGG GCC TCG 879
Gly Arg Gly Gly Leu Gly Ser Ile Phe Val Trp Ala Ser
285 290

GGG AAC GGG GGG CGG GAA CAT GAC AGC TGC AAC TGC GAC 918
Gly Asn Gly Gly Arg Glu His Asp Ser Cys Asn Cys Asp
295 300 305

GGC TAC ACC AAC AGT ATC TAC ACG CTG TCC ATC AGC AGC 957
Gly Tyr Thr Asn Ser Ile Tyr Thr Leu Ser Ile Ser Ser
310 315

GCC ACG CAG TTT GGC AAC GTG CCG TGG TAC AGC GAG GCC 996
Ala Thr Gln Phe Gly Asn Val Pro Trp Tyr Ser Glu Ala
320 325 330

TGC TCG TCC ACA CTG GCC ACG ACC TAC AGC AGT GGC AAC 1035
Cys Ser Ser Thr Leu Ala Thr Thr Tyr Ser Ser Gly Asn
335 340 345

CAG AAT GAG AAG CAG ATC GTG ACG ACT GAC TTG CGG CAG 1074
Gln Asn Glu Lys Gln Ile Val Thr Thr Asp Leu Arg Gln
350 355

AAG TGC ACG GAG TCT CAC ACG GGC ACC TCA GCC TCT GCC 1113
Lys Cys Thr Glu Ser His Thr Gly Thr Ser Ala Ser Ala
360 365 370

CCC TTA GCA GCC GGC ATC ATT GCT CTC ACC CTG GAG GCC 1152
Pro Leu Ala Ala Gly Ile Ile Ala Leu Thr Leu Glu Ala
375 380

AAT AAG AAC CTC ACA TGG CGG GAC ATG CAA CAC CTG GTG 1191
Asn Lys Asn Leu Thr Trp Arg Asp Met Gln His Leu Val
385 390 395

GTA CAG ACC TCG AAG CCA GCC CAC CTC AAT GCC AAC GAC 1230
Val Gln Thr Ser Lys Pro Ala His Leu Asn Ala Asn Asp
400 405 410

TGG GCC ACC AAT GGT GTG GGG CGG AAA GTG AGC CAC TCA 1269
Trp Ala Thr Asn Gly Val Gly Arg Lys Val Ser His Ser
415 420

TAT	GGC	TAC	GGG	CTT	TTC	GAC	GCA	GGC	GCC	ATG	GTG	GCC	1308
Tyr	Gly	Tyr	Gly	Leu	Leu	Asp	Ala	Gly	Ala	Mét	Val	Ala	
425				430						435			
CTG	GCC	CAG	AAT	TGG	ACC	ACA	GTC	GCC	CCC	CAG	CGG	AAG	1347
Leu	Ala	Gln	Asn	Trp	Thr	Thr	Val	Ala	Pro	Gln	Arg	Lys	
				440				445					
TGC	ATC	ATC	GAC	ATC	CTC	ACC	GAG	CCC	AAA	GAC	ATC	GGG	1386
Cys	Ile	Ile	Asp	Ile	Leu	Thr	Glu	Pro	Lys	Asp	Ile	Gly	
450				455					460				
AAA	CGG	CTC	GAC	CTC	CGG	AAC	ACC	GTG	ACC	GCG	TCC	CTG	1425
Lys	Arg	Leu	Glu	Val	Arg	Lys	Thr	Val	Thr	Ala	Cys	Leu	
				465			470			475			
GGC	GAG	CCC	AAC	CAC	ATC	ACT	CGG	CTG	GAG	CAC	GCT	CAG	1464
Gly	Glu	Pro	Asn	His	Ile	Thr	Arg	Leu	Glu	His	Ala	Gln	
				480				485					
GCG	CGG	CTC	ACC	CTG	TCC	TAT	AAT	CGC	CGT	GGC	GAC	CTG	1503
Ala	Arg	Leu	Thr	Leu	Ser	Tyr	Asn	Arg	Arg	Gly	Asp	Leu	
				490		495				500			
GCC	ATC	CAC	CTG	GTC	AGC	CCC	ATG	GGC	ACC	CGC	TCC	ACC	1542
Ala	Ile	His	Leu	Val	Ser	Pro	Met	Gly	Thr	Arg	Ser	Thr	
				505			510						
CTG	CTG	GCA	GCC	AGG	CCA	CAT	GAC	TAC	TCC	GCA	GAT	GGG	1581
Leu	Leu	Ala	Ala	Arg	Pro	His	Asp	Tyr	Ser	Ala	Asp	Gly	
				515		520			525				
TTT	AAT	GAC	TGG	GCC	TTC	ATG	ACA	ACT	CAT	TCC	TGG	GAT	1620
Phe	Asn	Asp	Trp	Ala	Phe	Met	Thr	Thr	His	Ser	Trp	Asp	
				530		535				540			
GAC	GAT	CCC	TCT	GGG	GAG	TGG	GTC	CTA	GAG	ATT	GAA	AAC	1659
Glu	Asp	Pro	Ser	Gly	Glu	Trp	Val	Leu	Glu	Ile	Glu	Asn	
				545				550					
ACC	AGC	GAA	GCC	AAC	AAC	TAT	GGG	ACG	CTG	ACC	AAC	TCC	1698
Thr	Ser	Glu	Ala	Asn	Asn	Tyr	Gly	Thr	Leu	Thr	Lys	Phe	
				555		560				565			
ACC	CTC	GTA	CTC	TAT	GGC	ACC	GCC	CCT	GAC	GGG	CTC	CCC	1737
Thr	Leu	Val	Leu	Tyr	Gly	Thr	Ala	Pro	Glu	Gly	Leu	Pro	
				570			575						
GTA	CCT	CCA	GAA	AGC	AGT	GGC	TGC	AAG	ACC	CTC	ACG	TCC	1776
Val	Pro	Pro	Glu	Ser	Ser	Gly	Cys	Lys	Thr	Leu	Thr	Ser	
				580		585			590				
AGT	CAG	GCC	TGT	GTG	GTG	TGC	GAG	GAA	GGC	TTC	TCC	CTC	1815
Ser	Gln	Ala	Cys	Val	Val	Cys	Glu	Glu	Gly	Phe	Ser	Leu	

595

600

605

CAC CAG AAG AGC TGT GTC CAG CAC TGC CCT CCA GGC TTC 1854
 His Gln Lys Ser Cys Val Gln Asn Cys Pro Pro Gly Phe
 610 615

GCC CCC CAA GTC CTC GAT ACG CAC TAT AGC ACC GAG AAT 1893
 Ala Pro Gln Val Leu Asp Thr Asn Tyr Ser Thr Glu Asn
 620 625 630

GAC GTG GAG ACC ATC CGG GCC AGC GTC TGC GCC CCC TGC 1932
 Asp Val Glu Thr Ile Arg Ala Ser Val Cys Ala Pro Cys
 635 640

CAC GCC TCA TGT GCC ACA TGC CAG GGG CCG GCC CTG ACA 1971
 His Ala Ser Cys Ala Thr Cys Gln Gly Pro Ala Leu Thr
 645 650 655

GAC TGC CTC AGC TGC CCC AGC CAC GCC TCC TTG GAC CCT 2010
 Asp Cys Leu Ser Cys Pro Ser His Ala Ser Leu Asp Pro
 660 665 670

GTG GAG CAG ACT TGC TCC CGG CAA AGC CAG AGC AGC CGA 2049
 Val Glu Gln Thr Cys Ser Arg Gln Ser Gln Ser Ser Arg
 675 680

GAG TCC CCG CCA CAG CAG CAG CCA CCT CGG CTG CCC CCG 2088
 Glu Ser Pro Pro Gln Gln Gln Pro Pro Arg Leu Pro Pro
 685 690 695

GAG GTG GAG GCG GGG CAA CGG CTG CGG GCA GGG CTG CTG 2127
 Glu Val Glu Ala Gly Gln Arg Leu Arg Ala Gly Leu Leu
 700 705

CCC TCA CAC CTG CCT GAG GTG GTG GCC GGC CTC AGC TGC 2166
 Pro Ser His Leu Pro Glu Val Val Ala Gly Leu Ser Cys
 710 715 720

GCC TTC ATC GTG CTG GTC TTC GTC ACT GTC TTC CTG CTC 2205
 Ala Phe Ile Val Leu Val Phe Val Thr Val Phe Leu Val
 725 730 735

CTG CAG CTG CGC TCT GGC TTT AGT TTT CGG GGG GTG AAG 2244
 Leu Gln Leu Arg Ser Gly Phe Ser Phe Arg Gly Val Lys
 740 745

GTG TAC ACC ATG GAC CGT GGC CTC ATC TCC TAC AAG GGG 2283
 Val Tyr Thr Met Asp Arg Gly Leu Ile Ser Tyr Lys Gly
 750 755 760

CTG CCC CCT GAA GCC TGG CAG GAG GAG TGC CCG TCT GAC 2322
 Leu Pro Pro Glu Ala Trp Gln Glu Glu Cys Pro Ser Asp
 765 770

TCA GAA GAG GAC GAG GGC CGG GGC GAG AGG ACC GCC TTT 2361
Ser Glu Glu Asp Glu Gly Arg Gly Glu Arg Thr Ala Phe
775 780 785

ATC AAA GAC CAG AGC GCC CTC TGA 2385
Ile Lys Asp Gln Ser Ala Leu End
790

(3) INFORMATION FOR SEQ ID NO:2

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 794 amino acids

(B) TYPE: amino acids

(C) STRANDEDNESS: single

(D) Topology: unknown

(ii) MOLECULE TYPE: furin

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(v) PUBLICATION INFORMATION: A.M.W. van den Ouweland et al,
Nucl. Acids. Res., 18:664 (1990)

(vi) SEQUENCE DESCRIPTION: SEQ ID NO:2

Met Glu Leu Arg Pro Trp Leu
1 5

Leu Trp Val Val Ala Ala Thr Gly Thr Leu Val Leu Leu
10 15 20

Ala Ala Asp Ala Gln Gly Gln Lys Val Phe Thr Asn Thr
25 30

Trp Ala Val Arg Ile Pro Gly Gly Pro Ala Val Ala Asn
35 40 45

Ser Val Ala Arg Lys His Gly Phe Leu Asn Leu Gly Gln
50 55

Ile Phe Gly Asp Tyr Tyr His Phe Trp His Arg Gly Val
60 65 70

Thr Lys Arg Ser Leu Ser Pro His Arg Pro Arg His Ser
75 80 85

Ala Thr Gln Phe Gly Asn Val Pro Trp Tyr Ser Glu Ala
320 325 330

Cys Ser Ser Thr Leu Ala Thr Thr Tyr Ser Ser Gly Asn
335 340 345

Gln Asn Glu Lys Gln Ile Val Thr Thr Asp Leu Arg Gln
350 355

Lys Cys Thr Glu Ser His Thr Gly Thr Ser Ala Ser Ala
360 365 370

Pro Leu Ala Ala Gly Ile Ile Ala Leu Thr Leu Glu Ala
375 380

Asn Lys Asn Leu Thr Trp Arg Asp Met Gln His Leu Val
385 390 395

Val Gln Thr Ser Lys Pro Ala His Leu Asn Ala Asn Asp
400 405 410

Trp Ala Thr Asn Gly Val Gly Arg Lys Val Ser His Ser
415 420

Tyr Gly Tyr Gly Leu Leu Asp Ala Gly Ala Met Val Ala
425 430 435

Leu Ala Gln Asn Trp Thr Val Ala Pro Gln Arg Lys
440 445

Cys Ile Ile Asp Ile Leu Thr Glu Pro Lys Asp Ile Gly
450 455 460

Lys Arg Leu Glu Val Arg Lys Thr Val Thr Ala Cys Leu
465 470 475

Gly Glu Pro Asn His Ile Thr Arg Leu Glu His Ala Gln
480 485

Ala Arg Leu Thr Leu Ser Tyr Asn Arg Arg Gly Asp Leu
490 495 500

Ala Ile His Leu Val Ser Pro Met Gly Thr Arg Ser Thr
505 510

Leu Leu Ala Ala Arg Pro His Asp Tyr Ser Ala Asp Gly
515 520 525

Phe Asn Asp Trp Ala Phe Met Thr Thr His Ser Trp Asp
530 535 540

Glu Asp Pro Ser Gly Glu Trp Val Leu Glu Ile Glu Asn
545 550

Thr Ser Glu Ala Asn Asn Tyr Gly Thr Leu Thr Lys Phe
555 560 565

Thr Leu Val Leu Tyr Gly Thr Ala Pro Glu Gly Leu Pro
570 575

Val Pro Pro Glu Ser Ser Gly Cys Lys Thr Leu Thr Ser
580 585 590

Ser Gln Ala Cys Val Val Cys Glu Glu Gly Phe Ser Leu
595 600 605

His Gln Lys Ser Cys Val Gln Asn Cys Pro Pro Gly Phe
610 615

Ala Pro Gln Val Leu Asp Thr Asn Tyr Ser Thr Glu Asn
620 625 630

Asp Val Glu Thr Ile Arg Ala Ser Val Cys Ala Pro Cys
635 640

His Ala Ser Cys Ala Thr Cys Gln Gly Pro Ala Leu Thr
645 650 655

Asp Cys Leu Ser Cys Pro Ser His Ala Ser Leu Asp Pro
660 665 670

Val Glu Gln Thr Cys Ser Arg Gln Ser Gln Ser Ser Arg
675 680

Glu Ser Pro Pro Gln Gln Gln Pro Pro Arg Leu Pro Pro
685 690 695

Glu Val Glu Ala Gly Gln Arg Leu Arg Ala Gly Leu Leu
700 705

Pro Ser His Leu Pro Glu Val Val Ala Gly Leu Ser Cys
710 715 720

Ala Phe Ile Val Leu Val Phe Val Thr Val Phe Leu Val
725 730 735

Leu Gln Leu Arg Ser Gly Phe Ser Phe Arg Gly Val Lys
740 745

Val Tyr Thr Met Asp Arg Gly Leu Ile Ser Tyr Lys Gly
750 755 760

Leu Pro Pro Glu Ala Trp Gln Glu Glu Cys Pro Ser Asp
765 770

Ser Glu Glu Asp Glu Gly Arg Gly Glu Arg Thr Ala Phe
775 780 785

Ile Lys Asp Gln Ser Ala Leu End
790